99

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/506,600	09/03/2004	Chung Tae Kim	YPL-PT019	2853

3624 7590 11/22/2006

VOLPE AND KOENIG, P.C.
UNITED PLAZA, SUITE 1600
30 SOUTH 17TH STREET
PHILADELPHIA, PA 19103

EXAMINER	
VAUTROT, DENNIS L	
ART UNIT	PAPER NUMBER
2167	

DATE MAILED: 11/22/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/506,600	KIM, CHUNG TAE	
	Examiner	**Art Unit**	
	Dennis L. Vautrot	2167	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on 31 August 2006.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1-34 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1-34 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on 03 September 2004 is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Response to Amendment

1. The applicants' amendment, filed 31 August 2006, has been received, entered into the record and considered.

2. As a result of the amendment, claims 23, 31, and 33 are amended. Claims 1 – 34 are pending in the application.

Response to Arguments

3. Applicant's arguments with respect to claims 7, 12, 13, 15, 26, 30 – 34 have been considered but are moot in view of the new ground(s) of rejection.

4. Applicant's arguments filed with respect to claims 1 – 6 as well as the generic argument relating to "the remaining claims" have been fully considered but they are not persuasive. In response to applicant's argument that the references fail to show certain features of applicant's invention, it is noted that the features upon which applicant relies (i.e., "A user can select the amount or length of multimedia contents which are searched using the same keyword.") are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

Claim Objections

5. The objection to claim 23 is withdrawn based on the amendment.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

6. Claims 1 – 4 are rejected under 35 U.S.C. 102(e) as being anticipated by **Gibbon et al.** (US 6,714,909).

7. Regarding claim 1, **Gibbon et al.** teaches a method of constructing a multimedia database method, comprising:

(a) receiving a start point and an end point of each first semantic unit of multimedia data, which is a smallest unit for searching for multimedia data (see column 11, lines 16-17 "The blocks in both sets are all time stamped, m=n and ..." This starting point and ending points are determined based on the time stamp provided. And see column 10, lines 41-43 "The goal is to extract three classes of semantics: news stories, augmented stories (augmented by the introduction of the story by the anchor), and

news summary of the day." These divisions represent different semantic units of multimedia data.);

(b) receiving a keyword for each first semantic unit (see column 13, lines 46-49 "For textual representation, keywords are chosen in step 5080 above, from the story according to their importance computed as weighted frequency.");

(c) receiving a start point and an end point of each second semantic unit of the multimedia data including at least one first semantic unit (see column 11, lines 16-17 "The blocks in both sets are all time stamped, m=n and ..." The starting point and ending points are determined based on the time stamp provided. And see column 10, lines 41-43 "The goal is to extract three classes of semantics: news stories, augmented stories (augmented by the introduction of the story by the anchor), and news summary of the day." These divisions represent different semantic units of multimedia data. The second semantic stricture is represented by the augmented stories – it is made up of the first semantic unit plus the news anchor summary.); and

(d) storing a keyword together with location information of its corresponding first semantic unit and second semantic unit (see column 13, lines 27-29 "The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia database 380 in step 5090." And see column 13, lines 50-52 "In the table of contents generated by the content description generator shown in FIG. 13, next to each story listed, a set of 10 keywords are given." The table of contents includes the keywords in the reference.)

8. Regarding claim 2, **Gibbon et al.** teaches (e) receiving a start point and an end point of each third semantic unit including a predetermined number of second semantic units, (see column 11, lines 16-17 "The blocks in both sets are all time stamped, m=n and ..." This starting point and ending points are determined based on the time stamp provided. And see column 10, lines 41-43 "The goal is to extract three classes of semantics: news stories, augmented stories (augmented by the introduction of the story by the anchor), and news summary of the day." These divisions represent different semantic units of multimedia data. The third semantic unit is interpreted to be the news summary of the day, which would incorporate keywords from the first and second semantic units below.)

wherein In (d), a keyword is stored with location information of its corresponding third semantic unit (see column 13, lines 27-29 "The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia database 380 in step 5090." And see column 13, lines 50-52 "In the table of contents generated by the content description generator shown in FIG. 13, next to each story listed, a set of 10 keywords are given.")

9. Regarding claim 3, **Gibbon et al.** teaches (f) receiving titles of each first semantic unit and each second semantic unit, (See column 3, lines 59-60 "Using the extracted stories and summaries/introductions, topics can be detected and categorized." The categories here can represent a title for the semantic unit.)

wherein in (d), a keyword is stored with the titles of its corresponding first semantic unit and second semantic unit. (See column 13, lines 50-52 "In the table of contents generated by the content description generator 455 shown in FIG. 13, next to each story listed, a set of 10 keywords are given." The keywords could also be used as titles.)

10. Regarding claim 4, **Gibbon et al.** teaches a keyword is classified into one of predetermined categories and is stored together with its corresponding category in (d). (See column 12, lines 50-56 "On the left of the screen, different semantics are categorized in the form of a table of contents...It is in a familiar hierarchical fashion which indexes directly into the stamped media data.")

Claim Rejections - 35 USC § 103

11. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

12. Claim 5 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** as applied to claim 1 or 4 above, and further in view of **Liu et al.** (US 6,970,860). **Gibbon et al.** teaches a method substantially as claimed. **Gibbon et al**. fails to teach a keyword is classified into a person category, an object category, a time category, or a place category. However **Liu et al.** teaches a keyword is classified into a person

category, an object category, a time category, or a place category (See Fig. 3, section 306, where the list of categories includes "People" as one of the options.) It would have been obvious to one with ordinary skill in the art to include the particular categories of classification into the method as disclosed in **Liu et al.** with the method as disclosed in **Gibbon et al.** because these are common useful categories that most media can be broken up into. It is for this reason that one of ordinary skill in the art would have been motivated to include a keyword is classified into a person category, an object category, a time category, or a place category.

13. Claim 6 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** as applied to claim 1 above, and further in view of **Liu et al.** (US 6,970,860). **Gibbon et al.** teaches a method substantially as claimed. **Gibbon et al.** fails to teach (g) storing a predetermined keyword together with its similar keywords so that a semantic unit corresponding to the predetermined keyword and semantic units corresponding to its similar keywords can be searched for together when a search for the semantic unit corresponding to the predetermined keyword or any of its similar keywords is carried out. However, **Liu et al.** teaches storing a predetermined keyword together with its similar keywords so that a semantic unit corresponding to the predetermined keyword and semantic units corresponding to its similar keywords can be searched for together when a search for the semantic unit corresponding to the predetermined keyword or any of its similar keywords is carried out. (See column 4, line 65 – column 5, line 4 "The feature and semantic matcher utilizes a semantic network to

locate objects with similar keywords. The semantic network defines associations between the keywords and multimedia objects. Weights are assigned to the associations to indicate how relevant certain keywords are to the multimedia objects.") It would have been obvious to one with ordinary skill in the art to combine the method as disclosed in **Gibbon et al.** with that in **Liu et al**. by adding the keyword combination storing feature because by having related keywords stored together the system can recognize similar keywords and provide more accurate results for the user, where only being able to recognize specific words would not have been as accurate. It is for this reason that one of ordinary skill in the art would have been motivated to store a predetermined keyword together with its similar keywords so that a semantic unit corresponding to the predetermined keyword and semantic units corresponding to its similar keywords can be searched for together when a search for the semantic unit corresponding to the predetermined keyword or any of its similar keywords is carried out.

14. Claim 7 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** as applied to claim 1 above, and further in view of **Matsuzawa et al.** (hereinafter, **Matsuzawa**, US 6,085,185). **Gibbon et al.** teaches a method substantially as claimed.

Gibbon et al. does not explicitly disclose the length of each first semantic unit and the length of each second semantic unit are determined by a user who constructs the multimedia database.

However, **Matsuzawa** discloses the length [range] of each first semantic unit and the length [range] of each second semantic unit are determined by a user who constructs the multimedia database. (See column 6, lines 12 – 22 "When a user viewing and operating a medium (video, etc.) puts an annotation (comment information) to a specific range or convenience of later retrieval, the annotation object management table 70 is a table for managing discrimination information of a medium to which the annotation is put, a specific range (in-point time code value and out-point time code value) and annotation information...together as an annotation object. Specific ranges to which an annotation is put may be partially overlapped with each other." Here, the user is determining the length of the semantic units which is represented by the in-point time and out-point time code values. Also because the reference notes that the ranges can overlap, this works with the other reference as allowing the different semantic unit levels to hierarchically share keywords based on their related semantic units.)

It would have been obvious to one with ordinary skill in the art at the time of the invention to combine the teachings of Gibbon et al. with that of Matsuzawa because both of the references are related to retrieval and indexing methods for a multi-media database, and by including the length of the semantic units being user selectable as disclosed in **Matsuzawa**, the method is more robust, allowing for any number of divisions of the data stream, as the user requires. It is for this reason that one of ordinary skill in the art would have been motivated to include disclose the length of each first semantic unit and the length of each second semantic unit are determined by a user who constructs the multimedia database.

15. Claims 8 –10 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** (US 6,714,909) in view of **Benitez et al.** (US 6,941,325) and further in view of **Nelson et al.** (US 6,243,713).

16. Regarding claim 8, **Gibbon et al.** teaches a system for constructing a multimedia database, comprising: a multimedia database which stores multimedia data (See column 4, lines 29-33 "The output of the multimedia content integration and description generation unit is stored in database 380 which can be subsequently retrieved upon a request from a user at terminal 390 through search engine 370." Here, database 380 is the multimedia database.); a keyword database which stores keywords necessary for searching for the multimedia data (See column 13, lines 27-29 "The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia database 380 in step 5090." The table of contents, as well as the multimedia description, include the keywords.); location information of each first semantic unit of the multimedia data, which is a smallest unit for searching for multimedia data, and location information of each second semantic unit of the multimedia data, which includes at least one first semantic unit (see column 11, lines 16-17 "The blocks in both sets are all time stamped, m=n and ..." The starting point and ending points are determined based on the time stamp provided. And see column 10, lines 41-43 "The goal is to extract three classes of semantics: news stories, augmented stories (augmented by the introduction

of the story by the anchor), and news summary of the day." These divisions represent different semantic units of multimedia data.);

Gibbon et al. fails to teach an input unit which receives the location information of each first semantic unit, including a start point and an end point, the location information of each second semantic unit, including a start point and an end point, and the keywords; and a control unit which receives the location information of each first semantic unit, the location information of each second semantic unit, and the keywords from the input unit and stores the keywords in the keyword database together with their corresponding first and second semantic units' location information.

However **Benitez et al.** teaches an input unit which receives the location information of each first semantic unit, including a start point and an end point, the location information of each second semantic unit, including a start point and an end point, and the keywords (See column 7, lines 43-47 "The media descriptor block includes information describing the media attributes of a cluster. For example, the media descriptor block may inherit format information, storage requirements, file identification parameters, and file location information of the clusters." In other words, the media descriptor block, by inheriting the location information and the keywords, is receiving all of the above mentioned information.) It would have been obvious to one with ordinary skill in the art to include an input unit as described in **Benitez et al.** with the database as described in **Gibbon et al.** because the importation of location and keyword information is necessary

for the proper functioning of a database which is based on this information. There had to have been some way of acquiring the data. It is for this reason that one of ordinary skill in the art would have been motivated to include an input unit which receives the location information of each first semantic unit, including a start point and an end point, the location information of each second semantic unit, including a start point and an end point, and the keywords.

In addition, **Nelson et al.** teaches a control unit which receives the location information of each first semantic unit, the location information of each second semantic unit, and the keywords from the input unit and stores the keywords in the keyword database together with their corresponding first and second semantic units' location information (See column 2, lines 56-67 "Alternatively, each of elements of the image may be separately stored in the multimedia index, each with data identifying the document and the position of the image in the document. The audio data would be indexed by speech recognition words or phonemes, each of which is indexed to reflect the audio's position at the 100th character, and further optionally indexed to reflect their relative time offset in the recorded audio. Thus, a single compound document can be indexed with respect to any number of multimedia components (or portions thereof), with the multimedia index reflecting the position of the multimedia component or its portions within the document".) It would have been obvious to one with ordinary skill in the art to combine the control unit as described in **Nelson et al.** with the database as described in **Gibbon et al.** and the input unit as described in **Benitez et al.** because after the data has been input in

any system, it must be properly stored. It is for this reason that one of ordinary skill in the art would have been motivated to include a control unit which receives the location information of each first semantic unit, the location information of each second semantic unit, and the keywords from the input unit and stores the keywords in the keyword database together with their corresponding first and second semantic units' location information.

17. Regarding claim 9, **Gibbon et al.** additionally teaches the input unit receives titles of each first semantic unit and each second semantic unit, and the control unit stores the titles in the keyword database together with their corresponding keywords. (See column 3, lines 59-60 "Using the extracted stories and summaries/introductions, topics can be detected and categorized." The categories here can represent a title for the semantic unit. And see column 13, lines 50-52 "In the table of contents generated by the content description generator 455 shown in FIG. 13, next to each story listed, a set of 10 keywords are given." The keywords could also be used as titles.)

18. Regarding claim 10, **Gibbon et al.** additionally teaches the input unit receives predetermined categories into which the keywords are classified, and the controller stores the keywords with their corresponding category. (See column 12, lines 50-56 "On the left of the screen, different semantics are categorized in the form of a table of contents...It is in a familiar hierarchical fashion which indexes directly into the stamped media data." The categories are found within the table of contents.)

19. Claim 11 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** (US 6,714,909) in view of **Benitez et al.** (US 6,941,325) and further in view of **Nelson et al.** (US 6,243,713) as applied to claim 8 above, and further in view of **Liu et al.** (US 6,970,860). **Gibbon et al., Benitez et al.**, and **Nelson et al.** teach a method substantially as claimed. **Gibbon et al., Benitez et al.**, and **Nelson et al.** fail to teach the keyword database includes a similar keyword database where keywords having similar meanings or indicating the same thing are stored, and when a keyword is input via the input unit, the controller searches the similar keyword database for a keyword that matches the input keyword, and, if there is a search result, stores the input keyword in the keyword database together with its similar keywords obtained from the similar keyword database so that not only a semantic unit corresponding to the input keyword but also semantic units corresponding to its similar keywords can be searched for when a search for the semantic unit of the input keyword or any of its similar keywords is carried out. However, **Liu et al.** teaches the keyword database includes a similar keyword database where keywords having similar meanings or indicating the same thing are stored, and when a keyword is input via the input unit, the controller searches the similar keyword database for a keyword that matches the input keyword, and, if there is a search result, stores the input keyword in the keyword database together with its similar keywords obtained from the similar keyword database so that not only a semantic unit corresponding to the input keyword but also semantic units corresponding to its similar keywords can be searched for when a search for the semantic unit of the

input keyword or any of its similar keywords is carried out. (See column 4, line 65 – column 5, line 4 "The feature and semantic matcher utilizes a semantic network to locate objects with similar keywords. The semantic network defines associations between the keywords and multimedia objects. Weights are assigned to the associations to indicate how relevant certain keywords are to the multimedia objects." And see column 6, lines 52-54 "For text queries, the feature and semantic matcher has a semantic matcher to identify objects with associated keywords that match the keywords from the query.") It would have been obvious to one with ordinary skill in the art to combine the method as disclosed in **Gibbon et al.** with that in **Liu et al**. by adding the keyword combination storing and searching feature because by having related keywords stored together the system can recognize similar keywords and provide more accurate results for the user, where only being able to recognize specific words would not have been as accurate. It is for this reason that one of ordinary skill in the art would have been motivated to have the keyword database includes a similar keyword database where keywords having similar meanings or indicating the same thing are stored, and when a keyword is input via the input unit, the controller searches the similar keyword database for a keyword that matches the input keyword, and, if there is a search result, stores the input keyword in the keyword database together with its similar keywords obtained from the similar keyword database so that not only a semantic unit corresponding to the input keyword but also semantic units corresponding to its similar keywords can be searched for when a search for the semantic unit of the input keyword or any of its similar keywords is carried out.

20. Claims 12 and 13 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** in view of **Sull et al.** (hereinafter **Sull**, US 2002/0069218).

21. Regarding claim 12, **Gibbon et al.** teaches a method of constructing a multimedia database, comprising: (a) setting a length of each first semantic unit of multimedia data, which is a smallest unit for searching for multimedia data according to a user's input (see column 11, lines 16-17 "The blocks in both sets are all time stamped, m=n and ..." This starting point and ending points are determined based on the time stamp provided. And see column 10, lines 41-43 "The goal is to extract three classes of semantics: news stories, augmented stories (augmented by the introduction of the story by the anchor), and news summary of the day." These divisions represent different semantic units of multimedia data. The user subsequently selects which division length they prefer to view.); and

(b) extracting a keyword from each first semantic unit using a predetermined method (see column 13, lines 46-49 "For textual representation, keywords are chosen in step 5080 above, from the story according to their importance computed as weighted frequency."); and

(d) storing the extracted keyword with its corresponding first semantic unit and second semantic unit (see column 13, lines 27-29 "The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia database 380 in step 5090." And see column 13, lines 50-52 "In the table of contents

generated by the content description generator shown in FIG. 13, next to each story listed, a set of 10 keywords are given." The table of contents includes the keywords in the reference.)

Gibbon does not explicitly disclose (c) setting a length of each second semantic unit of the multimedia data including at least one first semantic unit according to the users input.

However, **Sull** discloses setting a length of each second semantic unit of the multimedia data including at least one first semantic unit according to the users input. (See page 2, paragraph [0017] "In the case of video content, the segment may be a single frame, a single shot consisting of successive frames, or a group of several successive shots." And see page 4, paragraph [0032] "The metadata segments can form a hierarchical structure where the larger segment contains the smaller segments." This shows the length being set based on what the user wants, as well as the various segments [aka semantic units] being at least part of the other segments.)

It would have been obvious to one with ordinary skill in the art at the time of the invention to combine the teachings of **Gibbon et al.** with that of **Sull** because both are related to indexing multi media content and by including the hierarchical semantic units as disclosed in **Sull**, the method is more efficient by inheriting the keywords provided by lower level segments. It is for this reason that one of ordinary skill in the art would have been motivated to include setting a length of each second semantic unit of the multimedia data including at least one first semantic unit according to the users input.

22. Regarding claim 13, the combination of **Gibbon et al.** and **Sull** teaches (b1) extracting voice data from the multimedia data using a predetermined speech recognition technique (See **Gibbon et al.** column 3, lines 43-45 "Text may be from closed caption provided by a media provider or generated by the automatic speech recognition engine."); and (b2) extracting a predetermined part of speech from the extracted voice data as a keyword (see **Gibbon et al.** column 13, lines 46-49 "For textual representation, keywords are chosen in step 5080 above, from the story according to their importance computed as weighted frequency.")

23. Claims 14 and 15 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** in view of **Sull** as applied to claim 12 above, and further in view of **Liu et al.** (US 6,970,860).

24. Regarding claim 14, **Gibbon et al.** and **Sull** teaches a method substantially as claimed. **Gibbon et al.** and **Sull** do not explicitly disclose (b3) receiving a first keyword and first keyword information; and (b4) extracting the first keyword as a keyword of a first semantic unit when the first semantic unit has the same keyword information as the received keyword information. However, **Liu et al.** teaches (b3) receiving a first keyword and first keyword information; and (b4) extracting the first keyword as a keyword of a first semantic unit when the first semantic unit has the same keyword information as the received keyword information. (See column 8, lines 35-36 "At block 56, the retrieval/annotation system receives an initial query submitted by a user via the

user interface." And lines 44-45 "At block 604, the query handler parses the user query to extract one or more keywords." And column 10 lines 42-44 "...For each keyword in the input query, check if any of them is not in the keyword database.") It would have been obvious to one with ordinary skill in the art combine the teachings of **Liu et al.** and with the method as described in **Gibbon et al.** and **Sull** because by including a way to process the keywords inputted by the user with the keywords of the semantic unit, a match can be found to return as many as possible results. It is for this reason that one of ordinary skill in the art would have been motivated to include receiving a first keyword and first keyword information; and (b4) extracting the first keyword as a keyword of a first semantic unit when the first semantic unit has the same keyword information as the received keyword information.

25. Regarding claim 15, the combination of **Gibbon et al., Sull,** and **Liu et al**. teaches teach the keyword information is voice, an image, or text. (See **Gibbon et al**. column 2, lines 9-20 "The method may include separating a multimedia data stream into audio, visual and text components, segmenting the audio, visual and text components based on semantic differences..., identifying a topic of the multimedia event using the segmented text and topic category models...")

26. Claims 16 and 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** (US 6,714,909) in view of **Sull** and in view of **Nelson et al.** (US 6,243,713).

27. Regarding claim 16, **Gibbon et al.** teaches a system for constructing a multimedia database, comprising:

a multimedia database which stores multimedia data (See column 4, lines 29-33 "The output of the multimedia content integration and description generation unit is stored in database 380 which can be subsequently retrieved upon a request from a user at terminal 390 through search engine 370." Here, database 380 is the multimedia database.);

a keyword database which stores keywords necessary for searching for the multimedia data (See column 13, lines 27-29 "The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia database 380 in step 5090." The table of contents, as well as the multimedia description, include the keywords.); location information of each first semantic unit of the multimedia data, which is a smallest unit for searching for multimedia data, (see column 11, lines 16-17 "The blocks in both sets are all time stamped, m=n and ..." The starting point and ending points are determined based on the time stamp provided. And see column 10, lines 41-43 "The goal is to extract three classes of semantics: news stories, augmented stories (augmented by the introduction of the story by the anchor), and news summary of the day." These divisions represent different semantic units of multimedia data.);

a keyword extraction unit which extracts keywords from the multimedia data using a predetermined method (See column 13, lines 27-29 "The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia

database 380 in step 5090." The table of contents, as well as the multimedia description, include the keywords.);

Gibbon et al. fails to teach location information of each second semantic unit of the multimedia data, which includes at least one first semantic unit; and a control unit which divides the multimedia data into first semantic units and second semantic units and stores keywords in the keyword database together with their corresponding first and second semantic units' location information.

However, **Sull** teaches location information of each second semantic unit of the multimedia data, which includes at least one first semantic unit (See page 2, paragraph [0017] "In the case of video content, the segment may be a single frame, a single shot consisting of successive frames, or a group of several successive shots." And see page 4, paragraph [0032] "The metadata segments can form a hierarchical structure where the larger segment contains the smaller segments." This shows the length being set based on what the user wants, as well as the various segments [aka semantic units] being at least part of the other segments.) It would have been obvious to one with ordinary skill in the art at the time of the invention to combine the teachings of **Gibbon et al.** with that of **Sull** because both are related to indexing multi media content and by including the hierarchical semantic units as disclosed in **Sull**, the method is more efficient by inheriting the keywords provided by lower level segments. It is for this reason that one of ordinary skill in the art would have been motivated to include location information of each second semantic unit of the multimedia data, which includes at least one first semantic unit.

Also, **Nelson et al.** teaches a control unit which divides the multimedia data into first semantic units and second semantic units and stores keywords in the keyword database together with their corresponding first and second semantic units' location information. (See column 2, lines 56-67 "Alternatively, each of elements of the image may be separately stored in the multimedia index, each with data identifying the document and the position of the image in the document. The audio data would be indexed by speech recognition words or phonemes, each of which is indexed to reflect the audio's position at the 100^{th} character, and further optionally indexed to reflect their relative time offset in the recorded audio. Thus, a single compound document can be indexed with respect to any number of multimedia components (or portions thereof), with the multimedia index reflecting the position of the multimedia component or its portions within the document".) It would have been obvious to one with ordinary skill in the art to combine the control unit as described in **Nelson et al.** with the database as described in **Gibbon et al.** and **Sull** in order to properly store inputted data. It is for this reason that one of ordinary skill in the art would have been motivated to include a control unit which divides the multimedia data into first semantic units and second semantic units and stores keywords in the keyword database together with their corresponding first and second semantic units' location information.

28. Regarding claim 18, the combination of **Gibbon et al.**, **Sull**, and **Nelson et al.** additionally teaches the keyword extraction unit comprises: a voice extractor which extracts voice data from the multimedia data using a predetermined speech recognition

technique (See **Gibbon et al.** column 3, lines 43-45 "Text may be from closed caption provided by a media provider or generated by the automatic speech recognition engine."); and a part-of-speech extractor which extracts a predetermined part of speech from the voice data extracted by the voice extractor as a keyword (see **Gibbon et al.** column 13, lines 46-49 "For textual representation, keywords are chosen in step 5080 above, from the story according to their importance computed as weighted frequency.")

29. Claim 17 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** in view of **Sull** and in view of **Liu et al.** (US 6,970,860) as applied to claim 14 above and further in view of **Benitez et al.** (US 6,941,325). **Gibbon et al., Sull,** and **Liu et al.** teach a system substantially as claimed. **Gibbon et al., Sull,** and **Liu et al.** do not explicitly disclose an input unit which receives the location information of each first semantic unit, including a start point and an end point, the location information of each second semantic unit, including a start point and an end point, and the keywords. However **Benitez et al.** teaches an input unit which receives the location information of each first semantic unit, including a start point and an end point, the location information of each second semantic unit, including a start point and an end point, and the keywords. (See column 7, lines 43-47 "The media descriptor block includes information describing the media attributes of a cluster. For example, the media descriptor block may inherit format information, storage requirements, file identification parameters, and file location information of the clusters." In other words, the media descriptor block, by inheriting the location information and the keywords, is receiving all of the above

mentioned information.) It would have been obvious to one with ordinary skill in the art to include an input unit as described in **Benitez et al.** with the database as described in **Gibbon et al., Sull,** and **Liu et al.** because the importation of location and keyword information is necessary for the proper functioning of a database which is based on this information. There had to have been some way of acquiring the data. It is for this reason that one of ordinary skill in the art would have been motivated to include an input unit which receives the location information of each first semantic unit, including a start point and an end point, the location information of each second semantic unit, including a start point and an end point, and the keywords.

30. Claim 19 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** in view of **Sull** and in view of **Nelson et al.** as applied to claim 16 above, and further in view of **Liu et al.** (US 6,970,860). **Gibbon et al., Sull**, and **Nelson et al.** teach a system substantially as claimed. **Gibbon et al. Sull**, and **Nelson et al.** do not explicitly disclose an input unit which receives a first keyword and first keyword information, wherein the keyword extraction unit extracts the first keyword as a keyword of a first semantic unit when the first semantic unit has the same keyword information as the received keyword information. However, **Liu et al.** teaches an input unit which receives a first keyword and first keyword information, wherein the keyword extraction unit extracts the first keyword as a keyword of a first semantic unit when the first semantic unit has the same keyword information as the received keyword information. (See column 8, lines 35-36 "At block 56, the retrieval/annotation system receives an

initial query submitted by a user via the user interface." And lines 44-45 "At block 604, the query handler parses the user query to extract one or more keywords." And column 10 lines 42-44 "...For each keyword in the input query, check if any of them is not in the keyword database.") It would have been obvious to one with ordinary skill in the art combine the teachings of **Liu et al.** with the system as described in **Gibbon et al., Sull,** and **Nelson et al.** because by including a way to process the keywords inputted by the user with the keywords of the semantic unit, a match can be found to return as many as possible results. It is for this reason that one of ordinary skill in the art would have been motivated to include an input unit which receives a first keyword and first keyword information, wherein the keyword extraction unit extracts the first keyword as a keyword of a first semantic unit when the first semantic unit has the same keyword information as the received keyword information.

31. Claims 20-22 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** (US 6,714,909), in view of **Benitez et al.** (US 6,941,325), and further in view of **Liu et al.** (US 6,970,860).

32. Regarding claim 20, **Gibbon et al.** teaches a method of providing a multimedia data search service using a system for providing a multimedia data search service, including a multimedia database which stores multimedia data (See column 4, lines 29-33 "The output of the multimedia content integration and description generation unit is stored in database 380 which can be subsequently retrieved upon a request from a user

at terminal 390 through search engine 370." Here, database 380 is the multimedia database.); and a keyword database which stores keywords necessary for searching for the multimedia data (See column 13, lines 27-29 "The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia database 380 in step 5090." The table of contents, as well as the multimedia description, include the keywords.); location information of each first semantic unit of the multimedia data, which is a smallest unit for searching for multimedia data, and location information of each second semantic unit of the multimedia data, which includes at least one first semantic unit (see column 11, lines 16-17 "The blocks in both sets are all time stamped, m=n and ..." The starting point and ending points are determined based on the time stamp provided. And see column 10, lines 41-43 "The goal is to extract three classes of semantics: news stories, augmented stories (augmented by the introduction of the story by the anchor), and news summary of the day." These divisions represent different semantic units of multimedia data, some of which include the other – augmented new stories include new stories.),

(b) allowing a user to select a search unit level from between a first semantic unit and a second semantic unit (See column 4, lines 7-15 "The news data is segmented into multiple layers in a hierarchy to meet different needs. For instance, some users may want to retrieve a story directly; some others may want to listen to the news summary of the day in order to decide which story sounds interesting before making further choices...." Here, the different semantic units are represented by various lengths

of the data stream representing portions of a news broadcast. The user selects which length is the appropriate one for their use.); and

(d) outputting information of a searched semantic unit of the received search unit level, linking with the search semantic unit in the multimedia database (See column 4, lines 15-18 "This segmentation mechanism partitions the broadcast data in different ways so that direct indices to the events of different interests can be automatically established." In this case, the search unit level is the particular type of division of the broadcast and the index that is created creates the linking.)

Gibbon et al. fails to teach the method comprising: (a) receiving keywords necessary to search for multimedia data; (c) searching for multimedia data of the received search unit level whose keywords match the received keyword.

However **Benitez et al.** teaches receiving keywords necessary to search for multimedia data (See column 7, lines 43-47 "The media descriptor block includes information describing the media attributes of a cluster. For example, the media descriptor block may inherit format information, storage requirements, file identification parameters, and file location information of the clusters." In other words, the media descriptor block, by inheriting the location information and the keywords, is receiving all of the above mentioned information including the keywords.) It would have been obvious to one with ordinary skill in the art to include receiving the keywords as described in **Benitez et al.** with the database as described in **Gibbon et al.** because the importation of the keyword information is necessary for the proper functioning of a database which is based on this information. There had to have been some way of

acquiring the data. It is for this reason that one of ordinary skill in the art would have been motivated to include receiving keywords necessary to search for multimedia data.

Additionally, **Liu et al.** teaches searching for multimedia data of the received search unit level whose keywords match the received keyword. (See column 4, line 65 – column 5, line 4 "The feature and semantic matcher utilizes a semantic network to locate objects with similar keywords. The semantic network defines associations between the keywords and multimedia objects. Weights are assigned to the associations to indicate how relevant certain keywords are to the multimedia objects.") It would have been obvious to one with ordinary skill in the art to combine the method as disclosed in **Gibbon et al.** with that in **Liu et al**. by adding the keyword searching feature because by having related keywords stored together the system can recognize similar keywords and provide more accurate results for the user when searching for the keywords, whereas only being able to recognize specific words would not have been as accurate. It is for this reason that one of ordinary skill in the art would have been motivated to include searching for multimedia data of the received search unit level whose keywords match the received keyword.

33. Regarding claim 21, **Gibbon et al.** additionally teaches keywords are stored in the keyword database together with their corresponding first and second semantic units' location information and titles, and in (d), titles of searched semantic units are displayed on a screen. (See column 3, lines 59-60 "Using the extracted stories and summaries/introductions, topics can be detected and categorized." The categories here

can represent keywords or a title for the semantic unit. And see column 13, lines 50-52 "In the table of contents generated by the content description generator 455 shown in FIG. 13, next to each story listed, a set of 10 keywords are given." The keywords could also be used as titles.)

34. Regarding claim 22, **Gibbon et al**. additionally teaches the searched semantic units are displayed on a screen together with their respective keywords. (See column 13, lines 50-52 "In the table of contents generated by the content description generator 455 shown in FIG. 13, next to each story listed, a set of 10 keywords are given." The table of contents lists the searched semantic units.)

35. Claim 23 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** (US 6,714,909) in view of **Liu et al.** (US 6,970,860). **Gibbon et al.** teaches a system for providing a multimedia data search service, comprising: a multimedia database which stores multimedia data (See column 4, lines 29-33 "The output of the multimedia content integration and description generation unit is stored in database 380 which can be subsequently retrieved upon a request from a user at terminal 390 through search engine 370." Here, database 380 is the multimedia database.);

a keyword database which stores keywords necessary for searching for the multimedia data (See column 13, lines 27-29 "The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia database 380 in step 5090." The table of contents, as well as the multimedia description, include the

keywords.), location information of each first semantic unit of the multimedia data, which is a smallest unit for searching for multimedia data, and location information of each second semantic unit of the multimedia data, which includes at least one first semantic unit (see column 11, lines 16-17 "The blocks in both sets are all time stamped, m=n and ..." The starting point and ending points are determined based on the time stamp provided. And see column 10, lines 41-43 "The goal is to extract three classes of semantics: news stories, augmented stories (augmented by the introduction of the story by the anchor), and news summary of the day." These divisions represent different semantic units of multimedia data, with stories being a first unit and augmented stories being a second, meeting the part of the claim "which includes at least one first semantic unit.");

an input unit which receives a keyword and a search unit level a user (see column 13, lines 46-49 "For textual representation, keywords are chosen in step 5080 above, from the story according to their importance computed as weighted frequency." And see column 4, lines 7-15 "The news data is segmented into multiple layers in a hierarchy to meet different needs. For instance, some users may want to retrieve a story directly; some others may want to listen to the news summary of the day in order to decide which story sounds interesting before making further choices...." Here, the different semantic units are represented by various lengths of the data stream representing portions of a news broadcast. The user selects which length is the appropriate one for their use and it is received by the input unit.);

provides links between resulting search results and places in the multimedia database where the search results are stored, (See column 4, lines 15-18 "This segmentation mechanism partitions the broadcast data in different ways so that direct indices to the events of different interests can be automatically established." In this case, the search unit level is the particular type of division of the broadcast and the indices that are created, creates the linking.); and outputs some of the search results selected by the user; and a display unit which displays the searched results obtained by the control unit (See column 13, lines 50-52 "In the table of contents generated by the content description generator 455 shown in FIG. 13, next to each story listed, a set of 10 keywords are given." The table of contents lists the searched semantic units. As shown in FIG. 13, the results are also displayed for the user along the left column.)

Gibbon et al. does not explicitly disclose a control unit which searches the keyword database for a keyword that matches the received keyword.

However, **Liu et al.** teaches a control unit which searches the keyword database for a keyword that matches the received keyword (See column 4, line 65 – column 5, line 4 "The feature and semantic matcher utilizes a semantic network to locate objects with similar keywords. The semantic network defines associations between the keywords and multimedia objects. Weights are assigned to the associations to indicate how relevant certain keywords are to the multimedia objects.") It would have been obvious to one with ordinary skill in the art to combine the method as disclosed in **Gibbon et al.** with that in **Liu et al.** by adding the keyword searching feature because by having related keywords stored together the system can recognize similar keywords

and provide more accurate results for the user when searching for the keywords, whereas only being able to recognize specific words would not have been as accurate. It is for this reason that one of ordinary skill in the art would have been motivated to include a control unit which searches the keyword database for a keyword that matches the received keyword.

36. Claim 24 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Benitez et al.** (US 6,941,325) in view of **Gibbon et al.** (US 6,714,909). **Benitez et al.** teaches a method of constructing a multimedia database, comprising: (a) receiving location information of each semantic unit of multimedia data, which is a smallest unit for searching for multimedia data; (b) receiving a keyword for each semantic unit; (See column 7, lines 43-47 "The media descriptor block includes information describing the media attributes of a cluster. For example, the media descriptor block may inherit format information, storage requirements, file identification parameters, and file location information of the clusters." In other words, the media descriptor block, by inheriting the location information and the keywords, is receiving all of the above mentioned information including the keywords.)

Benitez et al. fails to teach storing keywords together with their corresponding semantic unit' location information.

However **Gibbon et al**. teaches storing keywords together with their corresponding semantic unit' location information. (See column 3, lines 59-60 "Using the extracted stories and summaries/introductions, topics can be detected and categorized."

The categories here can represent keywords for the semantic unit. And see column 13, lines 50-52 “In the table of contents generated by the content description generator 455 shown in FIG. 13, next to each story listed, a set of 10 keywords are given.” The categorization stores the keyword with the location information.) It would have been obvious to one with ordinary skill in the art to combine the method of **Benitez et al**. with that of **Gibbon et al.** because the keywords can be used to make finding the various appropriate semantic units possible during a search. It also becomes more efficient by storing them together. It is for this reason that one of ordinary skill in the art would have been motivated to include storing keywords together with their corresponding semantic unit' location information.

37. Claim 25 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** (US 6,714,909) in view of **Benitez et al.** (US 6,941,325) and further in view of **Nelson et al.** (US 6,243,713). **Gibbon et al.** teaches a system for constructing a multimedia database, comprising: a multimedia database which stores multimedia data (See column 4, lines 29-33 “The output of the multimedia content integration and description generation unit is stored in database 380 which can be subsequently retrieved upon a request from a user at terminal 390 through search engine 370.” Here, database 380 is the multimedia database.);

a keyword database which stores keywords necessary for searching for the multimedia data (See column 13, lines 27-29 “The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia database 380 in

step 5090." The table of contents, as well as the multimedia description, include the keywords.); and

location information of each semantic unit, which is a smallest unit for searching for multimedia data (see column 11, lines 16-17 "The blocks in both sets are all time stamped, m=n and ..." The starting point and ending points are determined based on the time stamp provided. And see column 10, lines 41-43 "The goal is to extract three classes of semantics: news stories, augmented stories (augmented by the introduction of the story by the anchor), and news summary of the day." These divisions represent different semantic units of multimedia data.);

Gibbon et al. fails to teach an input unit which receives the location information of each semantic unit, including a start point and an end point, and the keywords; and a control unit which receives the location information of each semantic unit from the input unit and stores the keywords in the keyword database together with their corresponding semantic units' location information.

However **Benitez et al.** teaches an input unit which receives the location information of each first semantic unit, including a start point and an end point, and the keywords (See column 7, lines 43-47 "The media descriptor block includes information describing the media attributes of a cluster. For example, the media descriptor block may inherit format information, storage requirements, file identification parameters, and file location information of the clusters." In other words, the media descriptor block, by inheriting the location information and the keywords, is receiving all of the above mentioned information.) It would have been obvious to one with ordinary skill in the art

to include an input unit as described in **Benitez et al.** with the database as described in **Gibbon et al.** because the importation of location and keyword information is necessary for the proper functioning of a database which is based on this information. There had to have been some way of acquiring the data. It is for this reason that one of ordinary skill in the art would have been motivated to include an input unit which receives the location information of each first semantic unit, including a start point and an end point, and the keywords.

In addition, **Nelson et al.** teaches a control unit which receives the location information of each semantic unit from the input unit and stores the keywords in the keyword database together with their corresponding semantic units' location information (See column 2, lines 56-67 "Alternatively, each of elements of the image may be separately stored in the multimedia index, each with data identifying the document and the position of the image in the document. The audio data would be indexed by speech recognition words or phonemes, each of which is indexed to reflect the audio's position at the 100th character, and further optionally indexed to reflect their relative time offset in the recorded audio. Thus, a single compound document can be indexed with respect to any number of multimedia components (or portions thereof), with the multimedia index reflecting the position of the multimedia component or its portions within the document".) It would have been obvious to one with ordinary skill in the art to combine the control unit as described in **Nelson et al.** with the database as described in **Gibbon et al.** and the input unit as described in **Benitez et al.** because after the data has been input in any system, it must be properly stored. It is for this reason that one of ordinary skill in

the art would have been motivated to include a control unit which receives the location information of each semantic unit from the input unit and stores the keywords in the keyword database together with their corresponding semantic units' location information.

38. Claim 26 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Sull** in view of **Gibbon et al. Sull** teaches a method of constructing a multimedia database, comprising: (a) receiving a length of each semantic unit of multimedia data, which is a smallest unit for searching for multimedia data according to a user's input (See page 2, paragraph [0017] "In the case of video content, the segment may be a single frame, a single shot consisting of successive frames, or a group of several successive shots. Each segment may be described by some elementary semantic information using texts." Here, the user is able to pick which length the unit should be from a frame all the way up to successive shots.)

(b) extracting a keyword from each semantic unit of the multimedia data (see column 13, lines 46-49 "For textual representation, keywords are chosen in step 5080 above, from the story according to their importance computed as weighted frequency." (See page 4, paragraph [0032] "Metadata of a video segment contain textual information such as time information for example, starting frame number and duration, or stating frame number as well as the finishing frame number), title, keyword, and annotation, as well as image information such as the key frame of a segment.").

Sull does not explicitly disclose (c) storing keywords together with its corresponding semantic unit's location.

However, **Gibbon et al.** discloses (c) storing keywords together with its corresponding semantic unit's location (see column 13, lines 27-29 "The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia database 380 in step 5090." And see column 13, lines 50-52 "In the table of contents generated by the content description generator shown in FIG. 13, next to each story listed, a set of 10 keywords are given." The table of contents includes the keywords in the reference.)

It would have been obvious to one with ordinary skill in the art at the time of the invention to combine the teachings of **Gibbon et al.** with that of **Sull** because both are related to indexing multi media content and by including the keyword and location information as disclosed in **Gibbon et al.**, the method is able to efficiently location the content based on the keywords. It is for this reason that one of ordinary skill in the art would have been motivated to include storing keywords together with its corresponding semantic unit's location.

39. Claim 27 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** (US 6,714,909) in view **Nelson et al.** (US 6,243,713). **Gibbon et al.** teaches a system for constructing a multimedia database, comprising: a multimedia database which stores multimedia data (See column 4, lines 29-33 "The output of the multimedia content integration and description generation unit is stored in database 380 which can

be subsequently retrieved upon a request from a user at terminal 390 through search engine 370." Here, database 380 is the multimedia database.); a keyword database which stores keywords necessary for searching for the multimedia data (See column 13, lines 27-29 "The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia database 380 in step 5090." The table of contents, as well as the multimedia description, include the keywords.); and location information of each semantic unit, which is a smallest unit for searching for multimedia data (see column 11, lines 16-17 "The blocks in both sets are all time stamped, m=n and ..." The starting point and ending points are determined based on the time stamp provided. And see column 10, lines 41-43 "The goal is to extract three classes of semantics: news stories, augmented stories (augmented by the introduction of the story by the anchor), and news summary of the day." These divisions represent different semantic units of multimedia data.); a keyword extraction unit which extracts keywords from the multimedia data using a predetermined method (see column 13, lines 46-49 "For textual representation, keywords are chosen in step 5080 above, from the story according to their importance computed as weighted frequency.");

Gibbon et al. fails to teach a control unit which divides the multimedia data into semantic units having a predetermined length and stores the extracted keywords together with their corresponding semantic units location information.

However, **Nelson et al.** teaches a control unit which divides the multimedia data into semantic units having a predetermined length and stores the extracted keywords together with their corresponding semantic units location information. (See column 2,

lines 56-67 "Alternatively, each of elements of the image may be separately stored in the multimedia index, each with data identifying the document and the position of the image in the document. The audio data would be indexed by speech recognition words or phonemes, each of which is indexed to reflect the audio's position at the 100th character, and further optionally indexed to reflect their relative time offset in the recorded audio. Thus, a single compound document can be indexed with respect to any number of multimedia components (or portions thereof), with the multimedia index reflecting the position of the multimedia component or its portions within the document".) It would have been obvious to one with ordinary skill in the art to combine the control unit as described in **Nelson et al.** with the database as described in **Gibbon et al.** because after the data has been input in any system, it must be properly stored. It is for this reason that one of ordinary skill in the art would have been motivated to include a control unit which divides the multimedia data into semantic units having a predetermined length and stores the extracted keywords together with their corresponding semantic units location information.

40. Claim 28 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** (US 6,714,909) in view of **Liu et al.** (US 6,970,860). **Gibbon et al.** teaches a method for providing a multimedia data search service system Including a multimedia database which stores multimedia data (See column 4, lines 29-33 "The output of the multimedia content integration and description generation unit is stored in database 380

which can be subsequently retrieved upon a request from a user at terminal 390 through search engine 370." Here, database 380 is the multimedia database.); and a keyword database which stores keywords necessary for searching for the multimedia data (See column 13, lines 27-29 "The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia database 380 in step 5090." The table of contents, as well as the multimedia description, include the keywords.); and location information of each semantic unit, which is a smallest unit for searching for multimedia data, (see column 11, lines 16-17 "The blocks in both sets are all time stamped, m=n and ..." The starting point and ending points are determined based on the time stamp provided. And see column 10, lines 41-43 "The goal is to extract three classes of semantics: news stories, augmented stories (augmented by the introduction of the story by the anchor), and news summary of the day." These divisions represent different semantic units of multimedia data.); the method comprising:

(a) inputting a keyword for searching for multimedia data (see column 13, lines 46-49 "For textual representation, keywords are chosen in step 5080 above, from the story according to their importance computed as weighted frequency." And see column 4, lines 7-15 "The news data is segmented into multiple layers in a hierarchy to meet different needs. For instance, some users may want to retrieve a story directly; some others may want to listen to the news summary of the day in order to decide which story sounds interesting before making further choices....");

(c) linking resulting search results to their locations in the multimedia database where the search results are stored, (See column 4, lines 15-18 "This segmentation

mechanism partitions the broadcast data in different ways so that direct indices to the events of different interests can be automatically established." In this case, the search unit level is the particular type of division of the broadcast and the indices that are created creates the linking.); and presenting the search results to a user (See column 13, lines 50-52 "In the table of contents generated by the content description generator 455 shown in FIG. 13, next to each story listed, a set of 10 keywords are given." The table of contents lists the searched semantic units. As shown in FIG. 13, the results are also displayed for the user along the self-column.)

Gibbon et al. does not explicitly disclose (b) searching for a semantic unit of a selected search unit level having the same keyword as the input keyword;

However, **Liu et al.** teaches (b) searching for a semantic unit of a selected search unit level having the same keyword as the input keyword; (See column 4, line 65 – column 5, line 4 "The feature and semantic matcher utilizes a semantic network to locate objects with similar keywords. The semantic network defines associations between the keywords and multimedia objects. Weights are assigned to the associations to indicate how relevant certain keywords are to the multimedia objects.") It would have been obvious to one with ordinary skill in the art to combine the method as disclosed in **Gibbon et al.** with that in **Liu et al**. by adding the keyword searching feature because by having related keywords stored together the system can recognize similar keywords and provide more accurate results for the user when searching for the keywords, whereas only being able to recognize specific words would not have been as accurate. It is for this reason that one of ordinary skill in the art would have been

motivated to include searching for a semantic unit of a selected search unit level having the same keyword as the input keyword.

41. Claim 29 is rejected under 35 U.S.C. 103(a) as being unpatentable over **Gibbon et al.** (US 6,714,909) in view of **Benitez et al.** (US 6,941,325) and further in view of **Liu et al.** (US 6,970,860). **Gibbon et al.** teaches a system for providing a multimedia data search service, comprising: a multimedia database which stores multimedia data (See column 4, lines 29-33 "The output of the multimedia content integration and description generation unit is stored in database 380 which can be subsequently retrieved upon a request from a user at terminal 390 through search engine 370." Here, database 380 is the multimedia database.);

a keyword database which stores keywords necessary for searching for the multimedia data (See column 13, lines 27-29 "The segmented content and multimedia descriptions (including the table of contents), are stored in multimedia database 380 in step 5090." The table of contents, as well as the multimedia description, include the keywords.); and location information of each semantic unit, which is a smallest unit for searching for multimedia data (see column 11, lines 16-17 "The blocks in both sets are all time stamped, m=n and ..." The starting point and ending points are determined based on the time stamp provided. And see column 10, lines 41-43 "The goal is to extract three classes of semantics: news stories, augmented stories (augmented by the introduction of the story by the anchor), and news summary of the day." These divisions represent different semantic units of multimedia data.); and

a display unit which displays the searched results obtained by the control unit. (See column 13, lines 50-52 "In the table of contents generated by the content description generator 455 shown in FIG. 13, next to each story listed, a set of 10 keywords are given." The table of contents lists the searched semantic units. As shown in FIG. 13, the results are also displayed for the user along the left column.)

Gibbon et al. fails to teach an input unit which receives a keyword from a user; a control unit which searches the keyword database for a keyword that matches the received keyword and outputs resulting search results with links to their locations in the multimedia database.

However **Benitez et al.** teaches an input unit which receives a keyword from a user (See column 12, lines 3-5 "The query processing subsystem can receive a user query through applicable input/output (I/O) circuitry..." The query makes up the keyword from the user.) It would have been obvious to one with ordinary skill in the art to include an input unit as described in **Benitez et al.** with the database as described in **Gibbon et al.** because the importation of the keyword is necessary for the proper functioning of a database which is based on or being searched using this information. There had to have been some way of acquiring the data. It is for this reason that one of ordinary skill in the art would have been motivated to include an input unit which receives a keyword from a user.

In addition, **Liu et al.** teaches a control unit which searches the keyword database for a keyword that matches the received keyword and outputs resulting search results with links to their locations in the multimedia database (See column 4, line 65 –

column 5, line 4 "The feature and semantic matcher utilizes a semantic network to locate objects with similar keywords. The semantic network defines associations between the keywords and multimedia objects. Weights are assigned to the associations to indicate how relevant certain keywords are to the multimedia objects.") It would have been obvious to one with ordinary skill in the art to combine the method as disclosed in **Gibbon et al.** with that in **Liu et al.** by adding the keyword searching feature because by having related keywords stored together the system can recognize similar keywords and provide more accurate results for the user when searching for the keywords, whereas only being able to recognize specific words would not have been as accurate. It is for this reason that one of ordinary skill in the art would have been motivated to include a control unit which searches the keyword database for a keyword that matches the received keyword and outputs resulting search results with links to their locations in the multimedia database.

42. Claims 30 – 34 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Milton** (US Patent Application Publication 2002/0059120) in view of **Matsuzawa**.

43. Regarding claim 30, **Milton** teaches a method of constructing a multimedia database, comprising:

(a) a user accessing a system (see page 2, paragraph [0023] "Thus a user can access his or her set of virtual inventory of media contents by simply using a web

enabled device at any web enabled location through his or her Media Access Provider.");

(b) allowing the user to designate address information of a multimedia data file desired to be executed by the user (see page 3, paragraph [0030] "The Virtual Content Handler serves as a resource for identifying the location of the media content associated with the Content Handle." Here, the user decides what multimedia data gets stored in the Content Handler, and the device keeps track of the location.);

(c) executing the multimedia data file by accessing a server where the multimedia data file is stored according to the designated address information (see page 6, paragraph [0073] "Once the request is authenticated, the media content owner streams the relevant media content directly to the user or via the media access provider of the user." In order to stream the content, the file has to be accessed from the server where it is stored.);

(d) receiving representative information of each first semantic unit (see page 4, paragraph [0042] "First, the content handle data element uniquely identifies a particular media content. Namely, the content handle is a universally recognized code that is assigned by a "virtual media registry" (VMR) to uniquely represent a particular media content, e.g., a particular CD of an artist, a particular video or movie and so on. This data element allows participating entitles within the Virtual Media Transactional Network to quickly associate the virtual inventory unit with a particular media content." By knowing the particular unit of the media, either a song, the whole cd, or however it is

broken up, the start and end location of the semantic unit would necessarily be included within the content handle data element.); and

(e) storing the representative information of each first semantic unit together with the start time and end time of each first semantic unit and the address information of the multimedia data file (see page 4, paragraph [0042] "Additionally, the content handle serves to describe the location as to where the virtual inventory units will be sent to be handled and rerouted. For example, the content handle is read by the VCH to determine the location of the media content to be accessed in the case of a 'content access request'".)

Milton does not explicitly disclose receiving and setting a start time and an end time of each first semantic unit of the multimedia data file, which is a smallest unit for searching for multimedia data while executing the multimedia data file. However **Matsuzawa** discloses receiving and setting a start time [in-point time code] and an end time [out-point time code] of each first semantic unit of the multimedia data file, which is a smallest unit for searching for multimedia data while executing the multimedia data file [viewing and operating a medium]. (See column 6, lines 12 – 22 "When a user viewing and operating a medium (video, etc.) puts an annotation (comment information) to a specific range for convenience of later retrieval, the annotation object management table 70 is a table for managing discrimination information of a medium to which the annotation is put, a specific range (in-point time code value and out-point time code value) and annotation information…together as an annotation object. Specific ranges to which an annotation is put may be partially overlapped with each other.")

It would have been obvious to one with ordinary skill in the art at the time of the invention to combine the teachings of **Milton** with that of **Matsuzawa** because both of the references are related to retrieval methods of a database containing multimedia, and by including the concurrently executing the data file teaching of **Matsuzawa**, the user is able to break the data apart in any number of different ways, providing for a more robust method. It is for this reason that one of ordinary skill in the art would have been motivated to include receiving and setting a start time and an end time of each first semantic unit of the multimedia data file, which is a smallest unit for searching for multimedia data while executing the multimedia data file.

44. Regarding claim 31, **Milton** teaches a system for constructing a multimedia database, comprising:

an input and output unit which allows a user to access a system (see page 2, paragraph [0023] "Thus a user can access his or her set of virtual inventory of media contents by simply using a web enabled device at any web enabled location through his or her Media Access Provider."), receives address information of a multimedia data file to be executed by the user (see page 3, paragraph [0030] "The Virtual Content Handler serves as a resource for identifying the location of the media content associated with the Content Handle." Here, the user decides what multimedia data gets stored in the Content Handler, and the device keeps track of the location.), a start time and an end time of each first semantic unit of the multimedia data file (see page 4, paragraph [0042] "First, the content handle data element uniquely identifies a particular media content.

Namely, the content handle is a universally recognized code that is assigned by a "virtual media registry" (VMR) to uniquely represent a particular media content, e.g., a particular CD of an artist, a particular video or movie and so on. This data element allows participating entitles within the Virtual Media Transactional Network to quickly associate the virtual inventory unit with a particular media content." By knowing the particular unit of the media, either a song, the whole cd, or however it is broken up, the start and end location of the semantic unit would necessarily be included within the content handle data element.), and representative information of each first semantic unit and allows the user to transmit data to or receive data from the server where the multimedia data file is stored (see page 4, paragraph [0042] "Additionally, the content handle serves to describe the location as to where the virtual inventory units will be sent to be handled and rerouted. For example, the content handle is read by the VCH to determine the location of the media content to be accessed in the case of a 'content access request'");

a keyword database which stores the representative information of each first semantic unit with the start time and end time of each first semantic unit and the address information of the multimedia data file (see page 4, paragraph [0042] "First, the content handle data element uniquely identifies a particular media content. Namely, the content handle is a universally recognized code that is assigned by a "virtual media registry" (VMR) to uniquely represent a particular media content, e.g., a particular CD of an artist, a particular video or movie and so on." This represents the representative information of each first semantic unit. And see page 4, paragraph [0042] "additionally,

the content handle serves to describe the location as to where the virtual inventory units will be sent to be handled and rerouted. For example, the content handle is read by the VCH to determine the location of the media content to be accessed in the case of a 'content access request'." This represents the address information.); and

a control unit which executes the multimedia data file by accessing the server where the multimedia data file is stored in response to an input from the user using the input and output unit (see page 6, paragraph [0073] "Once the request is authenticated, the media content owner streams the relevant media content directly to the user via the media access provider of the user."), and stores the received information in the keyword database together with the address information of the multimedia data file (see page 4, paragraph [0042] "Additionally, the content handle serves to describe the location as to where the virtual inventory units will be sent to be handled and rerouted. For example, the content handle is read by the VCH to determine the location of the media content to be accessed in the case of a 'content access request'." This represents the address information); and executes a predetermined first semantic unit of the multimedia data file using the address information of the multimedia data file and the start time and end time of the predetermined first semantic unit when a request for searching for and reproducing the predetermined first semantic unit is issued by the user (see page 6, paragraph [0074] "In step 330, method 300 plays the selected media content. In one illustrative embodiment, the media content owner forwards the stream of media content to a web enabled device specified by the user either directly or via the MAP of the user.")

Milton does not explicitly disclose receives the start time and end time while executing the multimedia file and the representative information of each first semantic unit in response to the input from the user. However, **Matsuzawa** discloses receives the start time [in-point time code] and end time [out-point time code] while executing the multimedia file and the representative information of each first semantic unit in response to the input from the user. (See column 6, lines 12 – 22 "When a user viewing and operating a medium (video, etc.) puts an annotation (comment information) to a specific range or convenience of later retrieval, the annotation object management table 70 is a table for managing discrimination information of a medium to which the annotation is put, a specific range (in-point time code value and out-point time code value) and annotation information…together as an annotation object. Specific ranges to which an annotation is put may be partially overlapped with each other."

It would have been obvious to one with ordinary skill in the art at the time of the invention to combine the teachings of **Milton** with that of **Matsuzawa** because both of the references are related to retrieval methods of a database containing multimedia, and by including the concurrently executing the data file teaching of **Matsuzawa**, the user is able to break the data apart in any number of different ways, providing for a more robust system. It is for this reason that one of ordinary skill in the art would have been motivated to include receives the start time and end time while executing the multimedia file and the representative information of each first semantic unit in response to the input from the user.

45. Regarding claim 32, **Milton** teaches a method of purchasing multimedia content from a multimedia content owner using a predetermined purchasing system, the method comprising:

(a) informing a user of purchasable multimedia contents and allowing the user to select multimedia content to be purchased (see page 3, paragraph [0034] "The Media Access Provider of the present invention provides the important function of creating and maintaining a virtual inventory of media contents.");

(b) executing the selected multimedia content using address information of the multimedia content stored in the purchasing system (see page 6, paragraph [0073] "Once the request is authenticated, the media content owner streams the relevant media content directly to the user or via the media access provider of the user.");

(d) storing the start time and end time of each first semantic unit of the multimedia content with the address information of the multimedia content (see page 4, paragraph [0042]); (see page 4, paragraph [0042] "Additionally, the content handle serves to describe the location as to where the virtual inventory units will be sent to be handled and rerouted. For example, the content handle is read by the VCH to determine the location of the media content to be accessed in the case of a 'content access request'." When the user picks which data content they would like to address, the start and end time necessarily go along with that b/c it would be stored in the content handle, which describes the content's location.);

(e) calculating a rate for a first semantic unit according to predetermined standards (see page 5, paragraph [0059]); and

(f) generating an execution file capable of executing a first semantic unit of the multimedia content purchased by the user using the start time and end time of the first semantic unit and the address information of the multimedia content stored in the purchasing system and providing information to which the execution file is linked (see page 6, paragraph [0074] "In step 330, method 300 plays the selected media content. In one illustrative embodiment, the media content owner forwards the stream of media content to a web enabled device specified by the user either directly or via the MAP of the user." The stream of media content would appear to be in the form of an execution file that is able to be used by the purchaser.)

Milton does not explicitly disclose (c) allowing the user to set a start time and an end time of each first semantic unit of the multimedia content, which is a smallest unit for purchasing the multimedia content, while executing the multimedia content. However, **Matsuzawa** discloses (c) allowing the user to set a start time and an end time of each first semantic unit of the multimedia content, which is a smallest unit for purchasing the multimedia content, while executing the multimedia content. (See column 6, lines 12 – 22 "When a user viewing and operating a medium (video, etc.) puts an annotation (comment information) to a specific range or convenience of later retrieval, the annotation object management table 70 is a table for managing discrimination information of a medium to which the annotation is put, a specific range (in-point time code value and out-point time code value) and annotation information…together as an annotation object. Specific ranges to which an annotation is put may be partially overlapped with each other."

It would have been obvious to one with ordinary skill in the art at the time of the invention to combine the teachings of **Milton** with that of **Matsuzawa** because both of the references are related to retrieval methods of a database containing multimedia, and by including the concurrently executing the data file teaching of **Matsuzawa**, the user is able to break the data apart in any number of different ways, providing for a more robust method. It is for this reason that one of ordinary skill in the art would have been motivated to include allowing the user to set a start time and an end time of each first semantic unit of the multimedia content, which is a smallest unit for purchasing the multimedia content, while executing the multimedia content.

46. Regarding claim 33, **Milton** teaches a system for purchasing multimedia content from a multimedia content owner using a predetermined purchasing system, the system comprising:

an input and output unit which allows a user to select multimedia content including a first semantic unit to be purchased and to set a start time and an end time of the first semantic unit (see page 4, paragraph [0042] "First, the content handle data element uniquely identifies a particular media content. Namely, the content handle is a universally recognized code that is assigned by a "virtual media registry" (VMR) to uniquely represent a particular media content, e.g., a particular CD of an artist, a particular video or movie and so on. This data element allows participating entitles within the Virtual Media Transactional Network to quickly associate the virtual inventory unit with a particular media content." By knowing the particular unit of the media, either

a song, the whole cd, or however it is broken up, the start and end location of the semantic unit would necessarily be included within the content handle data element.);

a keyword database which stores the start time and end time of the first semantic unit together with address information of multimedia contents that can be purchased using the purchasing system (see page 4, paragraph [0042] "Additionally, the content handle serves to describe the location as to where the virtual inventory units will be sent to be handled and rerouted. For example, the content handle is read by the VCH to determine the location of the media content to be accessed in the case of a 'content access request'." This represents the address information);

a controller which executes the selected multimedia content using the address information stored in the keyword database (see page 6, paragraph [0073] "Once the request is authenticated, the media content owner streams the relevant media content directly to the user via the media access provider of the user."); generates an execution file for executing the first semantic unit using the address information of the selected multimedia content and the start time and end time of the first semantic unit, and provides link information to which the execution file is linked (see page 6, paragraph [0074] "In step 330, method 300 plays the selected media content. In one illustrative embodiment, the media content owner forwards the stream of media content to a web enabled device specified by the user either directly or via the MAP of the user."); and

a rate calculation unit which calculates a rate for the first semantic unit according to predetermined standards (see page 5, paragraph [0049] "Second, the content doctrine can define or associate with a content handle a pricing hierarchy. For example,

a list of pricing for the media content can be associated with the type of transaction, e.g., wholesale, retail, promotion, and so on.")

Milton does not explicitly disclose stores the start time and end time of the first semantic unit in the keyword database in response to an input from the user using the input and output unit. However, **Matsuzawa** discloses stores the start time and end time of the first semantic unit in the keyword database in response to an input from the user using the input and output unit. (See column 6, lines 12 – 22 "When a user viewing and operating a medium (video, etc.) puts an annotation (comment information) to a specific range or convenience of later retrieval, the annotation object management table 70 is a table for managing discrimination information of a medium to which the annotation is put, a specific range (in-point time code value and out-point time code value) and annotation information...together as an annotation object. Specific ranges to which an annotation is put may be partially overlapped with each other."

It would have been obvious to one with ordinary skill in the art at the time of the invention to combine the teachings of **Milton** with that of **Matsuzawa** because both of the references are related to retrieval methods of a database containing multimedia, and by including the concurrently executing the data file teaching of **Matsuzawa**, the user is able to break the data apart in any number of different ways, providing for a more robust system. It is for this reason that one of ordinary skill in the art would have been motivated to include stores the start time and end time of the first semantic unit in the keyword database in response to an input from the user using the input and output unit.

47. Regarding claim 34, **Milton** teaches a computer-readable recording medium on which a program enabling the method of any of claims 1 through 4, 6, 7, 12 through 15, 20 through 22, 24, 26, 28, and 30 through 32 is recorded. (See page 4, paragraph [0040] "As such, the virtual content handler and the media content owner of the present invention can be stored on a computer readable medium.")

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dennis L. Vautrot whose telephone number is 571-272-2184. The examiner can normally be reached on Monday-Friday 9:00-6:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Cottingham can be reached on 571-272-7079. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Dv
13 November 2006

JOHN COTTINGHAM
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100